J. Mol. Biol. (1998) 277, 13–25

JMB

Characterization of ARC, a Divergent Member of the AAA ATPase Family from *Rhodococcus erythropolis*

Stefan Wolf[1], Istvan Nagy[2], Andrei Lupas[1], Günter Pfeifer[1] Zdenka Cejka[1], Shirley A. Müller[3], Andreas Engel[3], René De Mot[2] and Wolfgang Baumeister[1]*

[1]*Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-82152 Martinsried Germany*

[2]*F. A. Janssens Laboratory of Genetics, Catholic University of Leuven, B-3001 Heverlee Belgium*

[3]*Maurice E. Müller Institut Biozentrum, Universität Basel Klingelbergstr. 70 CH-4056 Basel, Switzerland*

**Corresponding author*

A gene encoding a AAA ATPase was discovered in the 5′ region of the second operon of 20 S proteasome subunits in the nocardioform actinomycete *Rhodococcus erythropolis* NI86/21. The gene was cloned and expressed in *Escherichia coli*. The protein, ARC (AAA ATPase forming Ring-shaped Complexes), is a divergent member of the AAA family. The deduced product of the *arc* gene is 591 residues long (66 kDa). The purified protein possesses a low, *N*-ethylmaleimide-sensitive ATPase activity and forms rings of six subunits, arranged symmetrically around a central opening or cavity. Two-dimensional crystals grown on lipid monolayers yielded images of the ATPase molecules in "end-on" orientation at 1.9 nm resolution.

Keywords: AAA family; Clp; Hsp100; ATPase; electron microscopy; *Rhodococcus erythropolis*

Introduction

The AAA ATPase (ATPases Associated with various Activities: Erdmann *et al.*, 1991) are a large family of ATPases involved in a wide array of cellular processes and occuring in all organisms. Twenty-one members of this family are present in the genome of the yeast *Saccharomyces cerevisiae* alone. These include Sec18, which is part of the machinery that fuses transport vesicles to their target membranes (Eakle *et al.*, 1988; Wilson *et al.*, 1989), Cdc48, which mediates homotypic membrane fusion during the cisternal regrowth of Golgi stacks and ER membranes (Rabouille *et al.*, 1995; Latterich *et al.*, 1995), Yta10-12, Yme1 and 26 S proteasome subunits, which are involved in protein degradation (Arlt *et al.*, 1996; Leonhard *et al.*, 1996; Lupas *et al.*, 1995), and Pas1, which is essential for peroxisome biogenesis (Erdmann *et al.*, 1991). The hallmark of these proteins is a highly conserved domain of approximately 250 residues, which contains the Walker A and B motifs (Walker *et al.*, 1982) typical of P-loop ATPases as well as a second region of high sequence conservation, C-terminal to the ATP-binding motifs (Erdmann *et al.*, 1991; Tomoyasu *et al.*, 1993). Proteins of the AAA family may contain one or two ATPase domains. The function and mechanism of the AAA domain is still unknown; its association with proteases in the 26 S proteasome and in FtsH homologues indicates an involvement in the disassembly and/or unfolding of proteins. Sequence comparisons show that the AAA proteins belong to a large superfamily of ATPases (Lupas *et al.*, 1997b), many of which are also involved in proteolysis and in the disassembly of protein complexes (Langer & Neupert, 1996; Schirmer *et al.*, 1996; Gottesman *et al.*, 1997). These include the Clp ATPases, Lon, and the helicase subunit RuvB.

During the characterization of two operons of 20 S proteasome subunits in the nocardioform actinomycete *Rhodococcus erythropolis* strain NI86/21 (Tamura *et al.*, 1995), we discovered a gene encoding a member of the AAA family approximately 1.5 kb 5′ of the second proteasome operon. Here we describe the cloning, sequencing and expression of this gene. The purified protein has a low, *N*-ethylmaleimide (NEM)-sensitive, Mg^{2+}-dependent ATPase activity and forms homo-hexameric rings. Because of this property we named it

Present address: A. Lupas, SmithKline Beecham Pharmaceuticals UP1345, 1250 South Collegeville Road, Collegeville, PA 19426-0989, USA.

Abbreviations used: NEM, *N*-ethylmaleimide; ORF, open reading frame; IPTG, isopropyl-β-D-thiogalactopyranoside; TEM, transmission electron microscopy; STEM, scanning transmission EM.

0022–2836/98/110013–13 $25.00/0/mb971589

ARC, for AAA ATPase forming Ring-shaped Complexes. ARC is a divergent member of the AAA ATPases, containing two insertions in the conserved domain, which are not observed in other members of the family. Unlike most AAA proteins from eubacteria, ARC is not a FtsH homologue; it branches off close to the root of the AAA family dendrogram and is not clearly assignable to any of its main branches.

Results

Gene organization

The DNA sequences of the two unlinked operons, $prcB_1A_1$ and $prcB_2A_2$, encoding the subunits of the 20 S proteasome of *R. erythropolis* were determined by Tamura *et al.* (1995). In both DNA regions, an open reading frame (*orf6*$_1$ and *orf6*$_2$, respectively, with the same orientation as the proteasome genes) was found 5′ of the proteasome operon. The same gene organization was identified by genomic sequencing in *Mycobacterium leprae* (Lupas *et al.*, 1997a) and, more recently, also in *Mycobacterium tuberculosis*, two other nocardioform actinomycetes. Further upstream, the two *Rhodococcus* DNA regions are quite different: the 5′ region of *orf6*$_1$ is occupied by the the operon, encoding the cytochrome P-450 system required for thiocarbamate degradation (Nagy *et al.*, 1995), whereas further sequence analysis now has shown that the 5′ region of *orf6*$_2$ is more similar to the equivalent DNA region in the mycobacteria. In the three species, a similar large open reading frame is present in the same orientation as the proteasome genes. However, the distances between these ORFs, encoding putative ATPases, and the downstream *orf6* genes is considerably smaller in *Rhodococcus* (about 1.5 kb) and *M. tuberculosis* (1.9 kb) as compared to *M. leprae* (3.8 kb). There is no obvious sequence similarity between the intergenic region of *Rhodococcus* and *Mycobacterium* (data not shown). The putative start codon of the *arc* gene in *Rhodococcus* is preceded by a ribosome binding site (GGGAGGAGA) at a distance of 5 bp. The gene is immediately followed by a potential stem-loop structure, which may be involved in transcription termination (stem length of 22 with three unpaired bases). A single RNA band of about 1.8 kb was visualized by Northern hybridization using an ATPase-specific probe (data not shown), consistent with a monocistronic mRNA terminated at the downstream hairpin structure.

Sequence analysis

The *arc* gene encodes a protein of 591 residues, which is 81% identical to the putative ATPases of *M. leprae* (GenBank U00017) and *M. tuberculosis* (GenBank Z97559) identified by genomic sequencing. From Ala11 to Gln77, it is predicted to contain a coiled-coil domain interrupted by Pro34 (Figure 1). In the mycobacterial homologues (which share 92% identical residues), this proline is embedded in an insertion of nine residues. The part of the coiled-coil following Pro34 has the same length and location within the sequence as the coiled-coil predicted in the 26 S proteasome ATPases; this region (Glu35 to Gln77) as well as the immediately adjacent sequence up to Glu147 can be aligned to the proteasome ATPases using a Gibbs sampling procedure (Lawrence *et al.*, 1993), but the significance of the scores is not sufficient to allow the assignment of this ATPase to the subfamily of proteasome ATPases. Although ARC clearly belongs to the AAA family of ATPases, it has no detectable sequence similarity to other AAA proteins outside the AAA domain.

The AAA domain of ARC extends from Val228 to Ile533 and contains all the sequence blocks typical for this family, as well as all the residues known or believed to be important for catalysis (Figure 2). It has a larger size due to two insertions of 15 and 33 residues (Ala293 to Ala307 and Val464 to Asn496, respectively), which are not seen in other AAA ATPases. In a distance-based dendrogram, it branches off close to the root of the AAA family and is not clearly assignable to any of its main branches (Figure 3). Sequence similarity is highest around the Walker A and B motifs, which are characteristic of the mononucleotide-binding fold (also known as the P-loop NTPase domain). This fold consists of a doubly wound, five-stranded, parallel β-sheet (for a structural classification of P-loop NTPases see the SCOP database at http://scop.mrc-lmb.cam.ac.uk/scop).

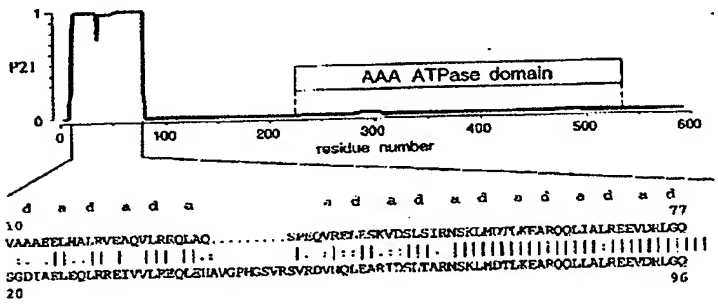

Figure 1. Domain structure of ARC. The AAA domain (Val228 to Ile533) is marked by an open bar. The graph shows the probability of coiled-coil formation calculated with a 21 residue scanning window. Below the graph, the regions containing the predicted coiled-coil segments in *Rhodococcus* and *Mycobacterium* ARC are aligned, and the residues forming the hydrophobic core of the heptad repeat are labeled a and d.

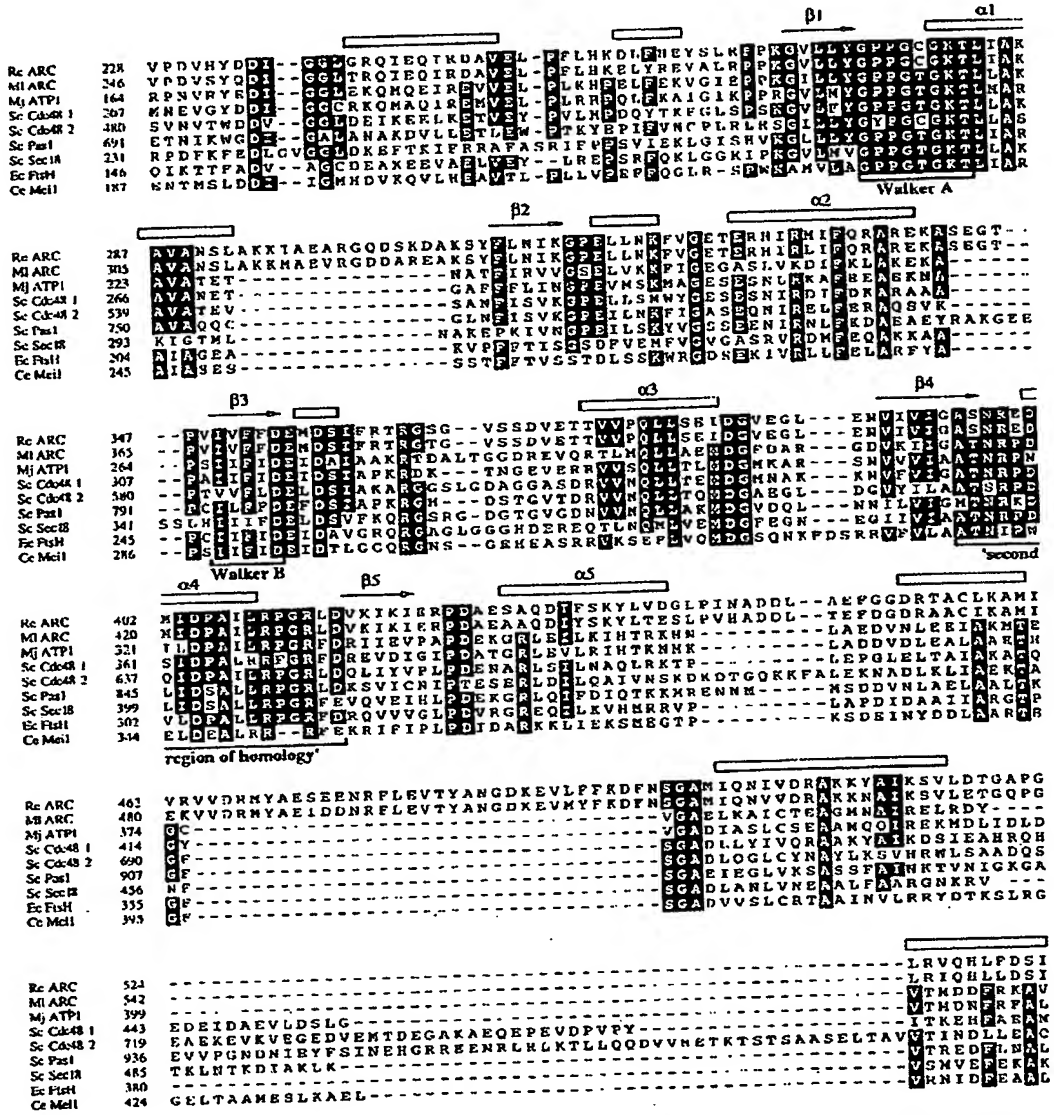

Figure 2. Alignment of *Rhodococcus* and *Mycobacterium* ARC with representative members of the AAA family. The organisms are: *Rhodococcus erythropolis* (Re); *Mycobacterium leprae* (Ml); *Methanococcus janaschii* (Mj); *Saccharomyces cerevisiae* (Sc); *Escherichia coli* (Ec); *Caenorhabditis elegans* (Ce). Residues conserved in a majority of sequences are shown in reverse type. The predicted consensus secondary structure is indicated by open boxes (α-helices) and arrows (β-strands). The secondary structure elements are numbered in the P-loop domain. The most highly conserved sequence regions corresponding to the Walker A and B domains (Walker *et al.*, 1982) and to the "second region of homology" (Tomoyasu *et al.*, 1993) are marked.

A consensus secondary structure prediction was performed in order to locate the structural elements of the nucleotide-binding fold within the conserved ATPase domain of the AAA family. This prediction represents the average of predictions by the method of Rost & Sander (1992), obtained separately for eight groups of AAA sequences: *Rhodococcus* and *Mycobacterium* ARC, 26 S proteasome ATPases, Cdc48 homologues (domain I), Cdc48 homologues (domain 2), ATPases involved in peroxisome biogenesis (domain 2), Sec18/NSF (domain 1), FtsH homol-

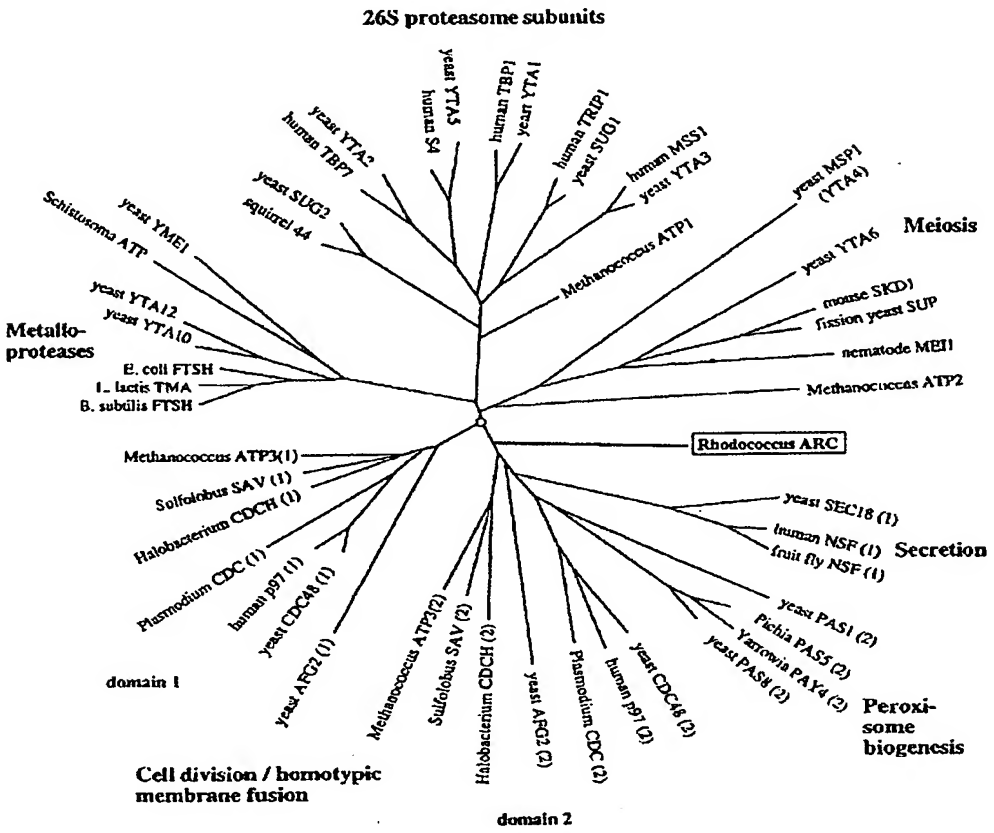

Figure 3. Distance-based dendrogram of the ATPase domains from proteins of the AAA family.

ogues, and Mei1 homologues. For the ATPases involved in peroxisomal biogenesis and Sec18/NSF, which have two ATPase domains, only one domain was included as the other is poorly conserved. These groups were selected on the basis of the dendrogram in Figure 3. Because of the large number of sequences considered and the wide range of pairwise sequence identities, the expected accuracy of the final consensus prediction is above 80% per residue. The canonical fold of P-loop NTPases was easily recognized in a central region containing five predicted β/α elements, which are numbered in Figure 2. This secondary structure assignment is validated by the known location of the Walker A and B motifs at the C-terminal ends of strands β1 and β3, respectively, and also by the amphipathic character of strands β2 and β5, which are predicted to be the outer strands of the central β-sheet and thus partly solvent-exposed. The central nucleotide-binding fold is flanked on either side by predicted α-helices; we refer to these regions as the helical extensions. Given that the N and C termini of the nucleotide-binding fold are in close proximity, it appears plausible that the helical extensions could form a common subdomain.

Heterologous expression in *E. coli* and purification

In order to obtain the enzyme in large quantities, we amplified the *arc* gene by PCR with primers encoding a C-terminally fused His_6-tag, and cloned the product into a phage T7 promoter-based expression vector. Induction with IPTG led to the production of a polypeptide of approximately 66 kDa as determined by SDS-PAGE (the calculated molecular mass of ARC is 66.1 kDa), with a yield of about 80 g per liter of cell culture. This polypeptide appeared only in cells carrying the recombinant plasmid (data not shown). Basal expression at a low level was detected in uninduced cells, possibly as a result of read-through from the T7 promoter. Because of the tendency of the protein to form inclusion bodies, the amount of soluble protein was increased by growing the cells at room temperature and inducing the expression with a reduced IPTG concentration.

The protein was purified by affinity chromatography on a Ni-NTA resin, from which it eluted in a single peak with approximately 80% purity. Fractionation by size exclusion chromatography on a Superose 6 column enhanced this purity to 95% (Figure 4(a)) and gave a yield of 20 to 30 mg per liter of cell culture. The N-terminal sequence was determined by Edman degradation and agreed with the deduced sequence of ARC but lacked the amino-terminal methionine. On the Superose 6 column, the protein behaved as a large oligomer of about 900 kDa. In a time-dependent manner, a minor second band appeared in the molecular mass range above 2 MDa (Figure 4(c)), indicating further aggregation of the oligomers. In native PAGE, the protein ran at about 769 kDa (Figure 4(b)), comparable to the 20 S proteasome of *R. erythropolis* NI86/21, indicating that under these conditions, the protein forms a dodecamer. The oligomeric ARC complexes were dissociated in 30% (w/v) urea, yielding one main dissociation product with the apparent molecular mass of an ARC dimer, but could not be reassembled after dialysis, whether or not the buffer contained ATP (data not shown).

In parallel to these experiments, a construct was generated carrying an N-terminally fused His_6-tag. In this form, the recombinant ARC protein could not be purified by affinity chromatography on a Ni-NTA resin, indicating that the N terminus of the protein is not accessible (at least in the oligomeric form).

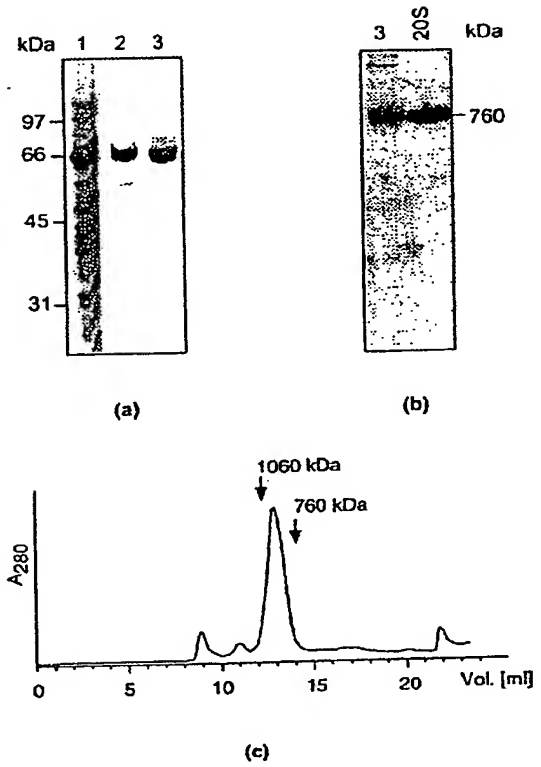

Figure 4. Gel electrophoresis of recombinant ARC complex stained with Coomassie blue after purification using different chromatographic steps. (a) Tricine SDS-10% PAGE show a dominant protein band of 66 kDa. Lane 1, membrane-free cell extract of expression host of BL-21 after IPTG induction; lane 2, pooled ARC containing fractions from Ni-NTA column; lane 3, Superose 6 main fraction. (b) Pooled high molecular mass fraction of Superose 6 was loaded on a native 5% to 15% gradient PAGE according to Laemmli (1970) at pH 8.8. The running position of the ATPase relative to the 20 S proteasome is indicated. (c) Typical Superose 6 chromatogram after running pooled Ni-NTA fractions containing ARC complexes. The main peak containing the ATPase elutes in a region between two internal standard proteins of molecular mass, 1060 kDa and 760 kDa for the thermosome from *T. acidophilum* and the 20 S proteasome from *R. erythropolis*, respectively.

Constitutive expression of the native ATPase protein

Polyclonal antibodies were raised against recombinant ARC in order to estimate its expression rate *in vivo*. To this end, the whole water-soluble cell extract of *R. erythropolis* NI86/21 was analyzed by two-dimensional gel electrophoresis, under denaturing conditions in the pH range of 4 to 6 (the calculated pI of ARC is 4.9 and the experimentally

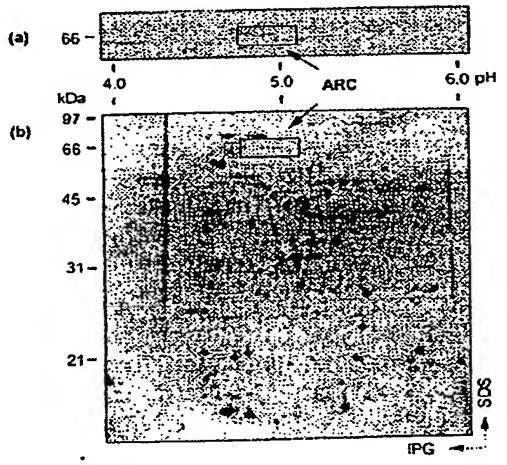

Figure 5. Two-dimensional electrophoresis of the crude cell extract of *R. erythropolis* from NI86/12. Protein samples (200 µg) in the first dimension were run on pH 4 to 6 IPG strips followed by SDS-PAGE in a 10% gel for running in the second dimension. After transfer onto nitrocellulose membranes, proteins were visualized by either polyclonal antiserum against recombinant ARC protein (a) or by Coomassie blue staining (b). The position of the ARC protein spots as judged by immunoblot analysis, is indicated with a box.

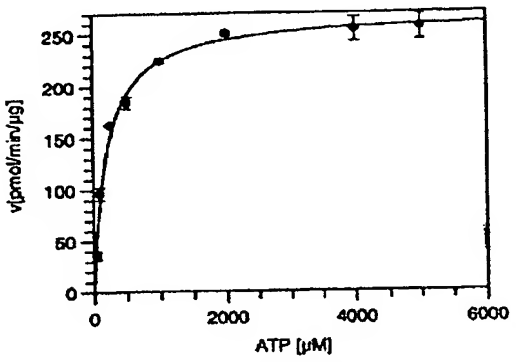

Figure 6. Plot of initial velocity *versus* ATP concentration.

Table 2. Effect of salt, inhibitor and nucleotide on the ATPase activity of recombinant ARC

Addition	Concentration (mM)	Activity (%)
Control		
$MgCl_2$	10	100
Salt		
EDTA	10	0
$CaCl_2$	10	0
Inhibitor		
NEM	10	5
NaN_3	3	100
ADP	1	100
Nucleotide (instead of ATP)		
ADP	1	5
GTP	1	0
CTP	1	50
UTP	1	0

Nucleotide hydrolyzing activity is expressed as a percentage of the control reaction, which contained 50 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 100 mM NaCl, 1 mM ATP.

determined pI of the recombinant protein is 5.2). Gels were run in parallel, blotted onto membranes, and either stained with Coomassie blue or incubated with the polyclonal antiserum. The location of native ARC in the two-dimensional pattern was determined by comparison of the two blots (Figure 5). ARC is expressed constitutively in a relatively small amount of less than 0.1 % of the total protein under normal growth conditions; this is significantly lower than the expression level of the 20 S proteasome in the same organism (Tamura *et al.*, 1995). Growth of cells at the maximum (34°C) and minimum (4°C) growth temperature did not cause a visible change of the expression rate. On immunostained membranes, ARC yielded a series of spots of the same molecular mass but of different pI values (Figure 5(a)), suggesting that the native protein may be multiply modified *in vivo*.

Nucleotide binding and hydrolysis

Recombinant ARC is purified as a stable oligomeric complex in the absence of ATP. The protein has a Mg^{2+}-dependent ATPase activity, whose kinetic parameters are comparable to those of other AAA protein (Figure 6 and Table 1). Preincubation with the sulfhydryl-blocking agent, *N*-ethylmaleimide (NEM, 10 mM) resulted in an almost complete inhibition of the ATPase activity, whereas the ATPase inhibitor sodium azide (NaN_3, 3 mM) had no effect. NEM sensitivity appears to be a general property of AAA ATPases, which is surprising given that no cysteine residue is highly conserved in this family. The pH optimum of ARC lies between pH 7 and 8 and drops off rapidly outside this range.

Nucleotide hydrolysis was also tested with ADP, GTP, CTP, and UTP (Table 2). CTP was cleaved at half the rate of ATP hydrolysis, but GTP or UTP were not substrates. ADP was cleaved at a low (5%) but detectable rate. No inhibition of ATPase activity by the reaction product ADP could be observed at equimolar concentration. This is in contrast to the observed inhibition of the AAA ATPase Cdc48, which is large even at substoichiometric amounts of ADP (Fröhlich *et al.*, 1995).

Table 1. Kinetic parameters of ATP hydrolysis by members of the AAA family

Protein	Organism	K_m (µM)	V_{max} (pmol/min per µg)	Reference
ARC	*Rhodococcus erythropolis*	200	268	This article
VAT	*Thermoplasma acidophilum*	n.d.	90	Pamnani *et al.* (1997)
CDC48	*Saccharomyces cerevisiae*	550	1560	Fröhlich *et al.* (1995)
p97	*Xenopus laevis*	n.d.	720	Peters *et al.* (1992)
NSF	Chinese hamster ovary	650	67	Tagaya *et al.* (1993)
	(high affinity site)	110	6	Morgan *et al.* (1994)
	(low affinity site)	43,300	257	Morgan *et al.* (1994)
VPS4p	*Saccharomyces cerevisiae*	n.d.	450	Babst *et al.* (1997)
Yta4p	*Saccharomyces cerevisiae*	5	7	Lucero *et al.* (1995)
SUG1	Rat liver	35	7	Makino *et al.* (1997)
26 S proteasome	Rabbit reticulocytes	15	19	Hoffman & Rechsteiner (1996)
RC	Rabbit reticulocytes	30	34	Hoffman & Rechsteiner (1996)

n.d., not determined

Structural analysis

The purified ARC protein was examined by negative stain electron microscopy (TEM) and by scanning transmission electron microscopy (STEM) of unstained/freeze-dried samples to obtain quantitive data regarding its mass. In both uranyl acetate and ammonium molybdate-stained preparations, the ring-shaped ARC protein is observed almost exclusively in an end-on orientation (Figure 7). The diameter of the complex is 10 to 11 nm, almost the same as the diameter of the 20 S proteasome. Image averaging clearly reveals 6-fold symmetry; the six subunits enclose a central heavily stain-filled opening, or cavity, traversing the complex (Figure 7(b)).

Some preliminary experiments undertaken were aimed at making two-dimensional crystals suitable for electron crystallography. We had some success with a lipid monolayer strategy using non-specific lipid mixtures from bovine brain extract, while experiments with Ni^{2+}NTA-chelator lipids (Schmitt *et al.*, 1994) designed to interact with the His_6-tag of ARC have not yet yielded useful crystals. Although coherent two-dimensional crystalline patches were rather small (Figure 8(a)) some image processing was possible. The resulting averages (Figure 8(b)) were in good agreement with the averages obtained from single particles but not superior to them. The power spectrum showed spots corresponding to a resolution of 1.92 nm; the resolution of the single particle average is 2.0 nm according to the radial correlation function criterion (Saxton & Baumeister, 1982).

STEM images of unstained freeze-dried ARC preparations showed a relatively non-homogeneous particle size distribution due to aggregation effects (Figure 9(a)). Nevertheless, evaluation of a data set of 1834 particles yielded a histogram with a single dominant peak at 368(±52) kDa ($n = 1661$) with some tailing towards masses up to three times this value (Figure 9(b)). On extrapolation to "zero dose" the measured mass becomes 384(±54) kDa, which is close to six times the molecular mass of the 66 kDa ARC monomer. In a control experiment, the ARC preparation was stabilized by glutaraldehyde crosslinking before adsorption. The mass histogram again displayed a single dominant peak, in this case at 415(±55) kDa ($n = 2897$) after extrapolation to zero dose (data not shown). A small population of particles had masses at approximately twice this value, 826(±96) kDa ($n = 90$). The 8% mass increase compared to the first result may be attributed to glutaraldehyde (Müller *et al.*, 1992). In gel filtration and in native PAGE we consistently found a higher molecular mass of approximately 760 kDa, indicating that under these conditions the ARC rings associate into pairs; a similar phenomenon has been observed with some other members of the AAA family (Hanson *et al.*, 1997; Pamnani *et al.*, 1997; Peters *et al.*, 1992) and with proteasomal α-rings (Zwickl *et al.*, 1994).

Discussion

Topology and active-site residues

ARC undoubtedly belongs to the AAA ATPases and is therefore built on a P-loop NTPase core, consisting of a doubly wound, five-stranded, parallel β-sheet. Among the NTPases of known structure containing this fold are nucleotide kinases, G proteins, myosin, kinesin, RecA, DEAD-box helicases, and F_1 ATPase. The nucleotide is bound at the C-terminal end of the central β-sheet in a cleft formed by loops connecting the central strands β1 and β3 to helices on opposite sides of the sheet.

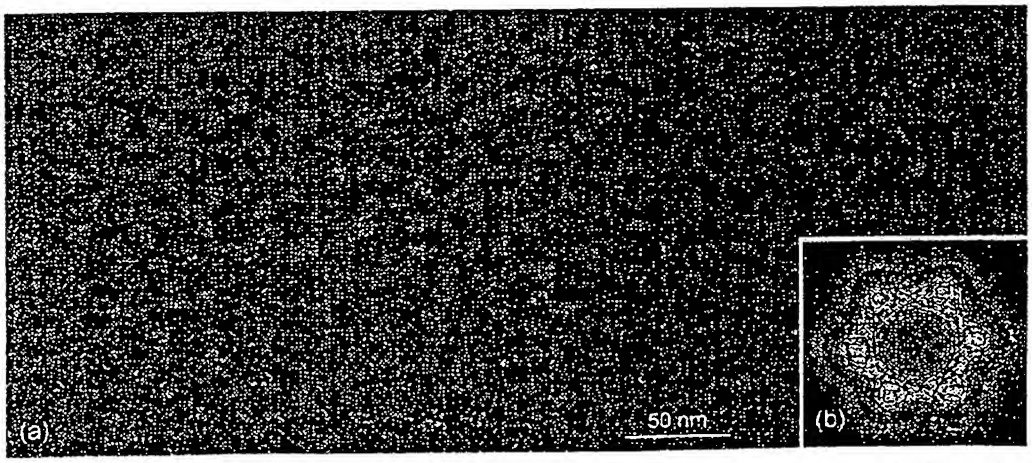

Figure 7. (a) Electron micrograph of purified recombinant ARC complexes negatively stained with 2% (w/v) uranyl acetate. Typically, images show individual ring-shaped molecules in their end-on orientation. (b) 6-fold symmetrized average of 1720 individual motifs. The diameter of the complex is 10 to 11 nm.

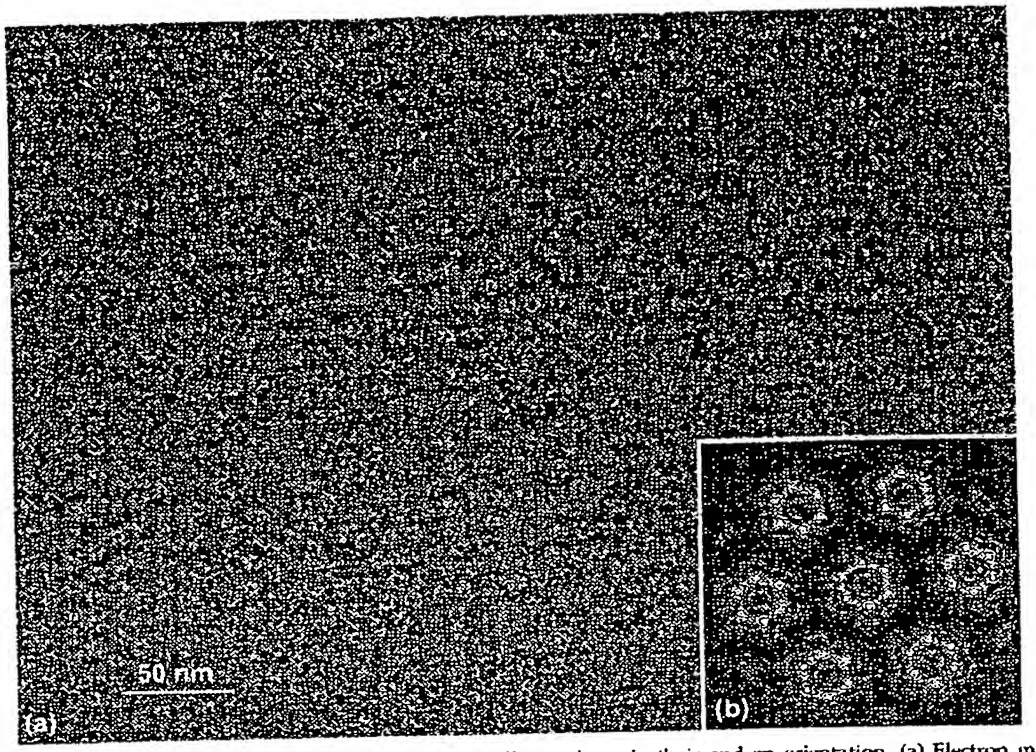

Figure 8. Two-dimensional crystal of recombinant ARC complexes in their end-on orientation. (a) Electron micrograph of a hexagonal two-dimensional crystalline array. (b) Correlation average derived from 410 individual motifs. The calculated diameter of the particle is around 10.5 nm.

These loops are the most highly conserved part of the fold: their sequences are known as the Walker A and B motifs (Walker *et al.*, 1982), and contain the residues essential for coordinating the phosphate groups of the nucleotide and the magnesium ion. In most P-loop NTPases, the Walker A and B motifs are adjacent, but in RecA, F_1 ATPase, and DEAD-box helicases, they are separated by strand β4, resulting in a different topology of the central β-sheet, although active-site residues occupy equivalent positions. It has been proposed that AAA proteins have a RecA-like topology and that, by analogy to RecA and F_1 ATPase, they use a glutamate at the C-terminal end of β2 for the nucleophilic attack (Yoshida & Amano, 1995). We note that this residue is not invariant in AAA proteins and poorly conserved in Clp ATPases and Lon (Lupas *et al.*, 1997b). In DEAD-box helicases (which also have a RecA-like topology) the glutamate of the DEAD motif at the C-terminal end of β3 initiates the nucleophilic attack (Subramanya *et al.*, 1996), and AAA, Clp, and Lon ATPases all contain a DExx motif after β3, suggesting that here also this glutamate is the nucleophilic residue.

In many P-loop NTPases, a residue C-terminal to β4 has been proposed to sense the presence of the γ-phosphate and transmit conformational changes resulting from nucleotide hydrolysis. This residue is Gln194 in RecA (Story & Steitz, 1992), Gln254 in the DExx-box helicase PrcA (Subramanya *et al.*, 1996), and Lys172 in 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (Hasemann *et al.*, 1996). In AAA proteins, two residues are likely candidates for this function: an asparagine at the C terminus of β4 (Asn398 in ARC), whose location corresponds to that of RecA Gln194 and PrcA Gln254, and an arginine at the N terminus of β5 (Arg412 in ARC), which is conserved in all AAA, Clp and Lon proteins (Lupas *et al.*, 1997b).

Is ARC a proteasome-like ATPase?

The *Rhodococcus* proteasome, like 20 S proteasomes from other sources, archaebacterial or eukaryotic, can only degrade unfolded polypeptides (Wenzel & Baumeister, 1995). Therefore, it requires the cooperation of an energy-dependent system binding the target proteins and presenting them in an unfolded form to the proteolytic 20 S core complex. Although there is conceptually no need for the formation of a stable protease-ATPase complex, a transient interaction seems sufficient, most known self-compartmentalizing proteases in fact form such complexes. While the evolutionary

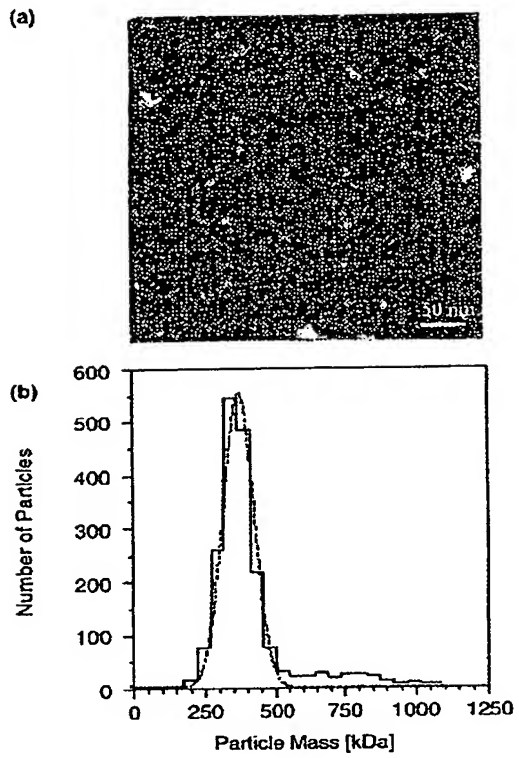

Figure 9. Mass determination of the ARC protein by means of STEM. (a) Elastic annular dark-field STEM images of purified recombinant ARC particles. (b) Histogram of particle/mass distribution of ARC shown with a fitted Gauss curve.

background of the protease components in such complexes is quite different, the ATPases are all P-loop NTPases with characteristic common sequence motifs beyond the Walker A and B motifs (Lupas *et al.*, 1997a). In eukaryotic proteasomes the ATPases are part of the 19 S cap complex which associates with the 20 S core to form the 26 S proteasomes. In addition to the six different but related ATPase subunits, the 19 S cap complex contains at least ten other subunits, some of which provide the link to the ubiquitin system. The proteasomal ATPases form a distinct subgroup of the AAA ATPase family (Figure 3). In archaebacteria a homologue of this subgroup has been identified (Bult *et al.*, 1996).

The occurrence of the *arc* gene in the vicinity of the proteasome operons in the actinomycetes *Rhodococcus* and *Mycobacterium*, and its absence from other eubacteria whose genome has been completely sequenced and which do not contain proteasomes raises the possibility that ARC functions in the context of the proteasome. Although the significance of blocks of sequence similarity to proteasome ATPases is too low to unambiguously classify it as a proteasomal ATPase, ARC has a similar topology, consisting of an N-terminal coiled-coil region, a central domain predicted to consist mainly of β-strands, and C-terminal AAA domain. This topology is typical for proteasome ATPases. It is interesting to note that recombinant ARC assembles spontaneously into a ring structure with almost exactly the dimensions of the 20 S proteasome. The two complexes differ, however, in symmetry, which is 7-fold in the case of the *Rhodococcus* proteasome (Zühl *et al.*, 1997) and 6-fold in the case of ARC. Such a mismatch begins to emerge as a common principle in protease-ATPase complexes as was noted for the ClpAP complex (Flanagan *et al.*, 1995; Kessel *et al.*, 1995) and might also apply to HslVU; in the latter case the isolated HslU complex shows some structural polymorphism, i.e. complexes built from six and seven subunits coexist (Rohrwild *et al.*, 1997). The functional significance of such a mismatch is still a matter for speculation.

So far we have not yet succeeded in reconstituting *in vitro* a functional ATPase-dependent system capable of degrading folded substrate proteins from purified *Rhodococcus* 20 S proteasomes and ARC. Thus, it is reasonable to assume that the association constant for the two protein complexes might be too low for detecting specific interactions outside the cell (Ellis, 1997). We further cannot rule out that additional factors are required for a successful reconstitution; the product of the highly conserved *orf7* gene, which in *Rhodococcus* and *Mycobacterium* is part of the proteasome gene operon (Tamura *et al.*, 1995; Knipfer & Shrader, 1997), is a possible candidate.

Materials and Methods

The nucleotide sequence of the *arc* gene was submitted to the GenBank with accession number AF028326.

All chemicals were of reagent grade or better and purchased as indicated. Solutions were prepared using Milli-Q water (Millipore).

Sequence analysis of ARC from *R. erythropolis* N186/21

Further characterization of the DNA region upstream of the proteasome structural genes $prcB_2$ and $prcA_2$, previously cloned in λ EMBL3 (clone FAJ2029: Tamura *et al.*, 1995), was carried out by DNA sequencing of overlapping subcloned fragments with an automated sequencer (A.L.F., Pharmacia Biotech). The sequence analysis of the region downstream of these proteasome genes has been reported (Nagy *et al.*, 1997b). Subsequent assembly was performed with the PCGENE software package (IntelliGenetics). For identification of potential coding regions, the GCWIND program (Shields *et al.*, 1992) was used.

Sequence similarity searches were performed using the BLAST WWW-page at NCBI (http://www.ncbi.nlm/nih.giov/cgi-bin/BLAST). The coiled-coil pre-

diction in Figure 1 was obtained from the COILS server at http://ulrec3.unil.ch/software/COILS_form.html. The alignment in Figure 2 was generated in MACAW (Schuler *et al.*, 1991). The dendrogram in Figure 3 was computed for the AAA ATPase domains only by the DARWIN E-mail server (cbrg@inf.ethz.ch). For the consensus secondary structure predictions, proteins from individual branches of the AAA family were aligned in PileUp (Genetics Computer Group, Madison, WI, USA), and the alignments were submitted to the PHD server (http://www.embl-heidelberg.de/predictprotein/phd_pred.html). After alignment of the entire family in MACAW, the predictions were averaged to yield the final prediction shown in Figure 2.

RNA isolation and Northern hybridization

Total RNA isolation from *R. erythropolis* NI86/21 and Northern analysis were carried out as described (Nagy *et al.*, 1997a). The riboprobe for Northern hybridization was prepared using pFAJ2555 linearized with *Ecl*136II as a template. This vector consists of pBlueScipt SK+ with the internal 468 bp *Bgl*II-*Pvu*II fragment of the ATPase gene cloned in the *Bam*HI-*Sma*I sites.

Cloning and overexpression experiments

For the enrichment of the ATPase gene by PCR, two sets of His_6-tag fused oligonucleotides based on the genomic N and C-terminal sequence were synthesized. The 5′ primers containing an *Nde*I site and the 3′ primers containing an *Eco*RI site had the following sequence: 5′-AT ATA TAT CAT ATG AGC TCG ACA GAG AAC CCG GAT TCG-3′ (5′AT ATA TAT CAT ATG CAT CAC CAT CAC CAT CAC ATG AGC TCG ACA GAG AAC CCG GAT T-3′) and 5′-TG GAA TTC CTA CTA CAG GTA CTG GCC CGT GTT CGA CTC-3′ (5′-TG GAA TTC CTA CTA GTG ATG GTG ATG GTG ATG CAG GTA CTG GCC CGT GTT CGA CTC-3′). PCR was performed with these primers and plasmid DNA (pFAJ2496) containing the complete ATPase gene using cloned *Pfu* polymerase (Stratagene) in a 25 µl reaction and cycled 25 times at 56°C annealing temperature. The PCR product was electrophoresed in 1.0% (w/v) agarose and the 1.8 kb band was excised from the gel and purified using the Qiaex DNA extraction kit (Qiagen). The eluted PCR fragment was cleaved with *Nde*I and *Eco* RI (Boehringer), ligated into dephosphorylated *Nde*I/*Eco*RI cut pT7-7 vector and transformed into host strain *E. coli* XL1-blue. Positive clones containing the ATPase gene were sequenced and transformed into the expression strain *E. coli* BL-21(DE3).

For screening for overexpression, small-scale cultures of selected clones were grown at 37°C in 3 ml of Luria broth (LB), containing ampicillin (100 µg/ml), to an A_{600} of 0.7, induced with 2.0 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG), and grown for a further four hours. The cells were harvested and the His_6-tagged protein purification was performed under native or denaturing conditions as described (QIAexpressionist, Qiagen). The level of induction was monitored using SDS-10% polyacrylamide gel electrophoresis. Under optimized conditions, the expression was performed in one litre of LB medium containing 100 µg/ml ampicillin inoculated 1:200 with the overnight culture and grown at room temperature until the cell density reached an A_{600} of 1.5, before 0.5 mM IPTG was added. The culture was then allowed to grow for a further two hours.

Purification of recombinant ARC protein from *E. coli*

The cells were harvested by centrifugation at 5000 *g* and lysed by sonication of a 20% (w/v) cell suspension in buffer A (50 mM sodium phosphate (pH 7.4), 300 mM NaCl, 1.5 mM β-mercaptoethanol) after incubation on ice for 30 minutes with 1 mg/ml lysozyme. After removal of the cell debris by centrifugation at 20,000 *g*, the supernatant was loaded with a flow rate of 15 cm/hour on a 4 ml Ni-NTA column, previously equilibrated with buffer A. The column was washed with the same buffer at pH 6.0 and bound proteins were eluted with a linear gradient of 0 to 0.5 mM imidazol in buffer A (10 vol/ml resin). Fractions containing the ARC were identified by SDS-10% PAGE. The semi-purified His_6-tagged protein was further purified on a Superose 6 column (Pharmacia), in buffer B (25 mM Tris-HCl (pH 7.4), 100 mM NaCl, 1 mM DTT) after it had been concentrated to 15 mg/ml using Centricon concentrators (Amicon). For storage, glycerol was added to the protein solution to a final concentration of 20% (w/v).

ATPase assays

Assays were routinely performed in 1.5 ml polypropylene microcentrifuge tubes for 30 minutes at 37°C. ATPase activity was determined by assaying the release of inorganic phosphate using the spectrophotometric method of Lanzetta *et al.* (1979) with modifications. A typical ATPase assay mixture (50 µl) contained 1.0 µg of freshly prepared ARC protein in buffer C (50 mM Tris-HCl (pH 7.4), 100 mM NaCl, 10 mM $MgCl_2$) and was started by the addition of 1 mM ATP. The reaction was stopped by the addition of 800 µl of reaction solution and 100 µl of 30% (w/v) citric acid and the absorption change was monitored using a photometer at a wavelength of 640 nm. To test for Mg^{2+} dependence the divalent cation was either chelated by the addition of 10 mM EDTA or replaced by 10 mM $CaCl_2$ in a Mg^{2+}-free buffer. When the nucleotide hydrolyzing activity was tested for other nucleoside phosphates, GTP, CTP, UTP and ADP instead of ATP were added at 1 mM initial concentration to start the reaction. The ATPase activity at various pH values (5.0, 6.0, 7.0, 8.0, 9.0) was determined in standard assay buffer C, where Tris-HCl was replaced by Tris-citrate of the appropriate pH. To test the sensitivity of the enzyme for ATPase inhibitors, *N*-ethylmaleimide (10 mM) or sodium azide (3 mM) were added to the assay mixture and incubated on ice for 15 minutes before the reaction was started. The kinetic parameters, K_m (Michaelis constant) and V_{max} (maximum velocity) were determined within the first ten minutes of assay conditions by measuring the initial velocity of ATP hydrolysis in triplicate. From the plot of the initial velocity of a simple Michaelis-Menten reaction *versus* the substrate concentration, the corresponding K_m and V_{max} values were calculated by non-linear regression methods (Cleland, 1979). It should be mentioned that the highest ATPase activity was always recorded with freshly prepared recombinant protein. The hydrolysis rate dropped down sharply within a few days when the enzyme was stored at 4°C. However, the addition of 20% glycerol had a stabilizing effect, which protected the activity almost completely.

Preparation of antiserum

Recombinant ARC protein (500 µg) purified by Ni-NTA and Superose 6 chromatography was dialyzed

against phosphate-buffered saline and homogenized in an equal volume of complete Freund's adjuvant (CFA). A New Zealand rabbit was immunized subcutaneously on day one followed by a second and third injection at days 21 and 42, respectively. Three weeks after the last injection whole blood was collected, allowed to clot for 60 minutes at 37°C and serum was obtained from the clot by removing the insoluble material by centrifugation.

Gel electrophoresis and immunoblotting

SDS-PAGE using 10% gels was performed as described (Schägger & von Jagow, 1987). Native gels with a linear gradient from 5% to 15% were run as described (Zühl *et al.*, 1997). Two-dimensional gel electrophoresis was performed using the method described by Görg *et al.* (1988). Proteins (200 µg) in the first dimension were run on immobilized pH gradient (IPG) gel strips with a linear gradient of pH 4 to 6 (prepared according to the description of the manufacturer; Pharmacia) and applied to denaturing SDS-10% PAGE for running in the second dimension. Visualization of the proteins was done after transfer onto nitrocellulose membranes according to Towbin *et al.* (1979) with either Coomassie-blue for total protein or with bromochloroindolyl phosphate/nitro blue tetrazolium (BCIP/NBT) for immunodetection with alkaline phosphatase.

Two-dimensional crystallization of ARC

Two-dimensional crystallization experiments were performed using a lipid monolayer technique. A small Teflon trough containing 500 µl of an aqueous subphase (10 mM Mes-KOH (pH 5.4), 200 mM NaCl, 20 mM $CdCl_2$, 5 mM $(NH_4)_2SO_4$, 2% (w/v) glucose) was used. Then 10 µl of a solution of 0.25 mg/ml bovine brain extract (Sigma) in chloroform/*n*-hexane (1:1, v/v) was spread onto it and allowed to stand for 15 minutes. The ARC protein solution (5 µl, 20 mg/ml) was then injected into the subphase, and incubated for 15 minutes to allow adsorption of the protein onto the lipid monolayer. The monolayer was transfered by placing a grid with a holey carbon film onto the surface and lifting it off. Samples were negatively stained with uranyl acetate (2% (w/v)) for transmission electron microscopy (TEM).

Transmission electron microscopy (TEM) and scanning transmission electron microscopy (STEM)

Samples from the Superose 6 fractions were brought to a concentration of 100 µg/ml and adsorbed onto glow-discharged carbon support films. TEM was performed using a Philips EM 420 at an accelerating voltage of 100 kV. Electron micrographs were routinely recorded at magnifications from 30,000× to 46,000×. Single particle averaging was performed as described (Baumeister *et al.*, 1988).

STEM mass measurement was performed as described by Engel (1978) to determine the molecular mass of the ARC protein. For this, the ATPase was diluted with 25 mM Mops buffer without β-mercaptoethanol, to a concentration of approximately 50 µg/ml. Adsorption was to glow-discharged thin carbon films supported by a fenestrated thick carbon layer on 200-mesh gold-coated copper grids. The grids were then washed four times on droplets of quartz-bidistilled water and freeze dried at −80°C in the microscope overnight. In a second experiment the complex was cross-linked with 0.1% (v/v) glutaraldehyde by a 30 minute incubation at room temperature prior to adsorption. The mass evaluations were carried out as described by Müller *et al.* (1992).

Acknowledgements

We thank F. Lottspeich for help with peptide sequencing, M. Boicu for DNA sequencing and B. Wolpensinger for preparing the ARC STEM samples. These studies were supported by a grant from the Human Frontiers Science Program to W. B. and the Maurice E. Müller Foundation. R. De M. is a Senior Research Associate with the Fund for Scientific Research (Flanders, Belgium).

References

Arlt, H., Tauer, R., Feldmann, H., Neupert, W. & Langer, T. (1996). The YTA10-12 complex, an AAA protease with chaperone-like activity in the inner membrane of mitochondria. *Cell*, 85, 875–885.

Babst, M., Sato, T. K., Banta, L. M. & Emr, S. D. (1997). Endosomal transport function in yeast requires a novel AAA-type ATPase, Vps4p. *EMBO J.* 16, 1820–1831.

Baumeister, W., Dahlmann, B., Hegerl, R., Kopp, F., Kuehan, L. & Pfeifer, G. (1988). Electron microscopy and image analysis of the multicatalytic proteinase. *FEBS Letters*, 241, 239–245.

Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J. F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. L., Geoghagen, N. S. M., Weidman, J. F., Fuhrmann, J. L., Nguyen, D., Utterback, T. R., Kelley, J. M., Peterson, J. D., Sadow, P. W., Hanna, M. C., Cotton, M. D., Roberts, K. M., Hurst, M. A., Kaine, B. P., Borodovsky, M., Klenk, H.-P., Fraser, C. M., Smith, H. O., Woese, C. R. & Venter, J. C. (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science*, 273, 1058–1073.

Cleland, W. W. (1979). Statistical analysis of enzyme kinetic data. *Methods Enzymol.* 63, 103–138.

Eakle, K. A., Bernstein, M. & Emr, S. D. (1988). Characterization of a component of the yeast secretion machinery: identification of the SEC18 gene product. *Mol. Cell. Biol.* 8, 4098–4109.

Ellis, R. J. (1997). Molecular chaperons: avoiding the crowd. *Curr. Biol.* 7, R531–R533.

Engel, A. (1978). Molecular weight determination by scanning transmission electron microscopy. *Ultramicroscopy*, 3, 273–281.

Erdmann, R., Wiebel, F. F., Flessau, A., Rykta, J., Beyer, A., Fröhlich, K. U. & Kunau, W. H. (1991). PAS1, a yeast required for peroxisome biogenesis, encodes a member of a novel family of putative ATPases. *Cell*, 64, 499–510.

Flanagan, J. M., Wall, J. S., Capel, M. S., Schneider, D. K. & Shanklin, J. (1995). Scanning transmission electron microscopy and small-angle scattering provide evidence that native *Escherichia coli* Cdc is a tetradecamer with an axial pore. *Biochemistry*, 34, 10910–10917.

Fröhlich, K. U., Fries, H. W., Peters, J. M. & Mecke, D. (1995). The ATPase activity of purified Cdc48p from *Saccharomyces cerevisiae* shows complex dependence on ATP-, ADP-, and NADH-concentrations and is completely inhibited by NEM. *Biochim. Biophys. Acta*, 1253, 25–32.

Görg, A., Postel, W., Domscheit, A. & Gunther, S. (1988). Two-dimensional electrophoresis with immobilized pH gradients of leaf proteins from barley (*Hordeum vulgare*): method, reproducibility and genetic aspects. *Electrophoresis*, 9, 681–692.

Gottesman, S., Wickner, S. & Maurizi, M. R. (1997). Protein quality control: triage by chaperones and proteases. *Genes Dev.* 11, 815–823.

Hanson, P. I., Roth, R., Morisaki, H., Jahn, R. & Heuser, J. E. (1997). Structure and conformational changes in NSF and its membrane receptor complexes visualized by quick-freeze/deep-etch electron microscopy. *Cell*, 90, 523–535.

Hasemann, C. A., Istvan, E. S., Uyeda, K. & Deisenhofer, J. (1996). The crystal structure of the bifunctional enzyme 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase reveals distinct domain homologies. *Structure*, 4, 1017–1029.

Hoffman, L. & Rechsteiner, M. (1996). Nucleotidase activities of the 26 S proteasome and its regulatory complex. *J. Biol. Chem.* 271, 32538–32545.

Kessel, M., Maurizi, M. R., Kim, B., Kocsis, E., Trus, B. L., Singh, S. K. & Steven, A. C. (1995). Homology in structural organization between *E. coli* ClpAP protease and the eukaryotic 26 S proteasome. *J. Mol. Biol.* 250, 587–594.

Knipfer, N. & Shrader, T. E. (1997). Inactivation of the 20 S proteasome in *Mycobacterium smegmatis*. *Mol. Microbiol.* 25, 375–383.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, 227, 680–685.

Langer, T. & Neupert, W. (1996). Regulated protein degradation in mitochondria. *Experientia*, 52, 1069–1076.

Lanzetta, P. A., Alvarez, L. J., Reinach, P. S. & Candia, O. A. (1979). An improved assay for nanomole amounts of inorganic phosphate. *Anal. Biochem.* 100, 95–97.

Latterich, M., Fröchlich, K. U. & Schekman, R. (1995). Membrane fusion and the cell cycle: Cdc48p participates in the fusion of ER membranes. *Cell*, 92, 885–893.

Lawrence, C. E., Altschul, S. F., Boguski, M. S., Neuwald, A. F. & Wootton, J. C. (1993). Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. *Science*, 262, 208–214.

Leonhard, K., Herrmann, J. M., Stuart, R. A., Mannhaupt, G., Neupert, W. & Langer, T. (1996). AAA proteases with catalytic sites on opposite membrane surfaces comprise a proteolytic system for the ATP-dependent degradation of inner membrane proteins in mitochondria. *EMBO J.* 15, 4218–4229.

Lucero, H. A., Chojnicki, E. W., Mandiyan, S., Nelson, H. & Nelson, N. (1995). Cloning and expression of a yeast gene encoding a protein with ATPase activity and high identity to the subunit 4 of the human 26 S protease. *J. Biol. Chem.* 270, 9178–9184.

Lupas, A., Zwickl, P., Wenzel, T., Seemüller, E. & Baumeister, W. (1995). Structure and function of the 20 S proteasome and of its regulatory complexes. *Cold Spring Harbor Symp. Quant. Biol.* 60, 515–524.

Lupas, A., Zühl, F., Tamura, T., Wolf, S., Nagy, I., De Mot, R. & Baumeister, W. (1997a). Eubacterial proteasomes. *Mol. Biol. Rep.* 24, 125–131.

Lupas, A., Flanagan, J. M., Tamura, T. & Baumeister, W. (1997b). Self-compartmentalizing proteases. *Trends Biochem. Sci.* 22, 399–404.

Makino, Y., Yamano, K., Kanemaki, M., Morikawa, K., Kishimoto, T., Shimbara, N., Tanaka, K. & Tamura, T. (1997). SUG1, a component of the 26 S proteasome, is an ATPase stimulated by specific RNAs. *J. Biol. Chem.* 272, 23201–23205.

Morgan, A., Dimaline, R. & Burgoyne, R. A. (1994). The ATPase activity of *N*-ethylmaleimide-sensitive fusion protein (NSF) is regulated by soluble NSF attachment proteins. *J. Biol. Chem.* 269, 29347–29350.

Müller, S. A., Goldie, K. N., Bürki, R., Häring, R. & Engel, A. (1992). Factors influencing the precision of quantitative scanning transmission electron microscopy. *Ultramicroscopy*, 46, 317–334.

Nagy, I., Schoofs, G., Compernolle, F., Proost, P., Vanderleyden, J. & de Mot, R. (1995). Degradation of the thiocarbamate herbicide EPTC (*S*-ethyl dipropylcarbamothioate) and biosafening by *Rhodococcus* sp. strain NI86/21 involve an inducible cytochrome P-450 system and aldehyde dehydrogenase. *J. Bacteriol.* 177, 676–687.

Nagy, I., Schoofs, G., De Schrijver, A., Vanderleyden, J. & De Mot, R. (1997a). New method for RNA isolation from actinomycetes: application to the transcriptional analysis of the alcohol oxidoreductase gene *thcE* in *Rhodococcus* and *Mycobacterium*. *Lett. Appl. Microbiol.* 25, 75–79.

Nagy, I., Schoofs, G., Vanderleyden, J. & De Mot, R. (1997b). Further sequence analysis of the DNA regions with the *Rhodococcus* 20 S proteasome structural genes reveals extensive homology with *Mycobacterium leprae*. *DNA Seq.* 7, 225–228.

Pamnani, V., Tamura, T., Lupas, A., Peters, J., Cejka, Z., Ashraf, W. & Baumeister, W. (1997). Cloning, sequencing and expression of VAT, a CDC48/p97 ATPase homologue from the archaeon *Thermoplasma acidophilum*. *FEBS Letters*, 404, 263–268.

Peters, J. M., Harris, J. R., Lustig, A., Müller, S., Engel, A., Volker, S. & Franke, W. W. (1992). Ubiquitous soluble Mg^{2+}-ATPase complex. A structural study. *J. Mol. Biol.* 223, 557–571.

Rabouille, C., Levine, T. P., Peters, J. M. & Warren, G. (1995). An NSF-like ATPase, p97, and NSF mediate cisternal regrowth from mitotic Golgi fragments. *Cell*, 82, 905–914.

Rohrwild, M., Pfeifer, G., Santarius, U., Müller, S. A., Huang, H. C., Engel, A., Baumeister, W. & Goldberg, A. L. (1997). The ATP-dependent HslVU protease from *Escherichia coli* is a four-ring structure resembling the proteasome. *Nature Struct. Biol.* 4, 133–139.

Rost, B. & Sander, C. (1992). Jury returns on structure prediction. *Nature*, 360, 540.

Saxton, W. O. & Baumeister, W. (1982). The correlation averaging of a regularly arranged bacterial cell envelope protein. *J. Microsc.* 127, 127–138.

Schägger, H. & von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* 166, 368–379.

Schirmer, E. C., Glover, J. R., Singer, M. A. & Lindquist, S. (1996). HSP100/Clp proteins: a common mechan-

ism explains diverse functions. *Trends Biochem. Sci.* 21, 289–296.
Schmitt, L., Dietrich, C. & Tampé, R. (1994). Synthesis and characterization of chelator-lipids for reversible immobilization of engineered proteins at self-assembled lipid interfaces. *J. Am. Chem. Soc.* 116, 8485–8491.
Schuler, G. D., Altschul, S. F. & Lipman, D. J. (1991). A workbench for multiple alignment construction and analysis. *Proteins: Struct. Funct. Genet.* 9, 180–190.
Shields, D. C., Higgins, D. G. & Sharp, P. M. (1992). GCWIND: a microcomputer program for identifying open reading frames according to codon positional G + C content. *Comput. Appl. Biosci.* 8, 521–532.
Story, R. M. & Steitz, T. A. (1992). Structure of the RecA protein-ADP complex. *Nature*, 355, 374–376.
Subramanya, H. S., Bird, L. E., Brannigan, J. A. & Wigley, D. B. (1996). Crystal structure of a DExx box DNA helicase. *Nature*, 384, 379–383.
Tagaya, M., Wilson, D. W., Brunner, M., Arango, N. & Rothman, J. E. (1993). Domain structure of an *N*-ethylmaleimide-sensitive fusion protein involved in vesicular transport. *J. Biol. Chem.* 268, 2662–2666.
Tamura, T., Nagy, I., Lupas, A., Lottspeich, F., Cejka, Z., Schoofs, G., Tanaka, K., De Mot, R. & Baumeister, W. (1995). The first characterization of a eubacterial proteasome: the 20 S complex of *Rhodococcus*. *Curr. Biol.* 5, 766–774.
Tomoyasu, T., Yuki, T., Morimura, S., Mori, H., Yamanaka, K., Niki, H., Hiraga, S. & Ogura, T. (1993). The *Escherichia coli* FtsH protein is a prokaryotic member of a protein family of putative ATPases involved in membrane function, cell cycle control, and gene expression. *J. Bacteriol.* 175, 1344–1351.
Towbin, H., Staehelin, T. & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl Acad. Sci. USA*, 76, 4350–4354.
Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1, 945–951.
Wenzel, T. & Baumeister, W. (1995). Conformational constraints in protein degradation by the 20 S proteasome. *Nature Struct. Biol.* 2, 199–204.
Wilson, D. W., Wilcox, C. A., Flynn, G. C., Chen, E., Kuang, W. J., Henzel, W. J., Block, M. R., Ullrich, A. & Rothman, J. E. (1989). A fusion protein required for vesicle-mediated transport in both mammalian cells and yeast. *Nature*, 339, 355–359.
Yoshida, M. & Amano, T. (1995). A common topology of proteins catalyzing ATP-triggered reactions. *FEBS Letters*, 359, 1–5.
Zühl, F., Tamura, T., Dolenc, I., Cejka, Z., Nagy, I., De Mot, R. & Baumeister, W. (1997). Subunit topology of the *Rhodococcus* proteasome. *FEBS Letters*, 400, 83–90.
Zwickl, P., Kleinz, J. & Baumeister, W. (1994). Critical elements in proteasome assembly. *Nature Struct. Biol.* 1, 765–770.

Edited by P. E. Wright

(Received 6 October 1997; received in revised form 12 December 1997; accepted 16 December 1997)